A Semantic Image Retrieval Framework Based on Ontology and Naïve Bayesian Inference

Roung Shiunn Wu[1], Wen Hsien Hsu[2]

[1, 2] Department of Information Management, National Chung Cheng University,

Chia Yi, Taiwan

[1]roungwu@ccu.edu.tw; [2]lanbosama@gmail.com

Abstract- **To deal with large amounts of multimedia content, the MPEG-7 standard was proposed in order to provide a general standard by which to describe various types of multimedia content. However, we required not only a standard to describe multimedia content but also a retrieval framework to search for the requested semantic content. The proposed framework employs ontology and MPEG-7 descriptors to deal with problems arising between syntactic representation and semantic retrieval of images. Instead of building a single ontology for a specific domain, the framework allows for the construction of incrementally multiple ontologies, and shares ontology information not only between the image seekers but also between different domains. Naïve Bayesian inference is used to estimate the similarity between query ontology and domain ontology for matching relevant images. The framework provides a relevant feedback mechanism for image seekers to respond to relevant images in the matching process, thus enabling the framework to enhance image annotations for improved image retrieval precision.**

Keywords- Semantic-based Image Retrieval; Ontology; Naïve Bayesian Classifier; Image Annotation; Relevance Feedback

I. INTRODUCTION

Advanced information technologies continue to be developed, so do more digital multimedia contents be produced and stored. To deal with multimedia content, the MPEG-7 standard was promoted as a common interface for audiovisual content description. The MPEG-7 standard provides a way to store and retrieve audiovisual information without the need to perform actual coding [11]. Most of the MPEG-7 semantic descriptors make it possible to embed information such as object, event, and location into digital contents.

However, MPEG-7 was defined as a generic schema to serve a variety of audiovisual applications, but not for use in specific retrieval applications. The effective exploitation of MPEG-7 in a retrieval procedure therefore remains an issue to be explored [15]. Due to its XML schema-based representation, MPEG-7 is not suitable for representing the semantic aspect of multimedia content in a formal and concise way. To consider syntactic and semantic information in the same framework, and to expand the function of MPEG-7 descriptors, applying domain ontology to extend the MPEG-7 standard is a promising approach, which can provide a shared and formal vocabulary for the specification of the image retrieval [20].

The diversity of multimedia data allows users to easily use multimedia content. However, the precise means of multimedia content retrieval for users becomes a critical issue. Two popular approaches to image retrieval are content-based and keyword-based retrievals. These approaches both have their advantages and disadvantages. A content-based approach requires advanced image processing and pattern recognition techniques, while a keyword-based approach needs image annotations, which could result in low retrieval precision. Thus, a semantic-based retrieval approach is proposed to address these problems. The aim is to develop a framework which deals with semantic image retrieval by using domain ontology. In our work, domain ontology is applied to semantic-based image retrieval with MPEG-7 descriptors. Although MPEG-7 descriptors have its disadvantages when it comes to the description of semantic image concepts, the power of the MPEG-7 standard is employed to describe low-level features. The domain ontology is used to compensate for this MPEG-7 limitation, narrowing the gap between syntactic and semantic concepts.

In order to improve semantic image retrieval performance, the proposed framework provides a mechanism which enhances image annotations during the inference stage. Due to the combination of both domain ontology and MPEG-7 descriptors, the framework provides not only semantic-based, but also content-based image retrieval. On one hand, the MPEG-7 descriptors are used to represent the syntactic content; on the other hand, domain ontology is used to describe high-level semantic concepts. In this way, with high-level semantic annotations [9], semantic image retrieval precision can be markedly improved.

The remainder of this paper is organized as follows: in Section II, we review related works; the proposed framework is described in Section III; more details of design and implementations are provided in Section IV; the experimental results are presented in Section V; and conclusions are provided in Section VI.

II. RELATED WORKS

This section reviews related works on the use of MPEG-7, ontology, and inference techniques in image retrieval. The proposed framework explores the use of MPEG-7 descriptors to retrieve images with low-level features, and ontology-based methods of image retrieval using semantic level context.

A. Multimedia Retrieval Based on MPEG-7

Most text-based search engines have difficulty in modelling and extracting the concept of relevance and summarization in multimedia retrieval applications, due to the richness of audiovisual content. It is thus difficult for users to query exact information from multimedia content. The Multimedia Content Description Interface (MPEG-7 [1]) was standardized to specify a rich set of tools for various types of multimedia information, facilitate the quick and efficient

[1]http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

identification of interesting and relevant information, and provide metadata for describing the features of multimedia content.

A framework of a multimedia retrieval system [11] based on MPEG-7 was proposed to provide a rich set of automatic feature extraction components and an independent retrieval interface. The database system proposed in [16] highlights the most relevant aspects considered during the design and implementation of a DBMS-driven MPEG-7 layer on top of a content-based multimedia retrieval system. An efficient method was proposed in [18] for compactly representing colour and texture features for image retrieval. The method used MPEG-7 visual descriptors for colour, and homogeneous texture description for texture representation.

B. Ontology-based Image Retrieval

Ontology is defined as an explicit specification of a conceptualization. It consists of several components, including: concepts, relationships, attributes, instances, and axioms. Ontology defines the semantic of concepts and their inter-relationships for a specific domain. It thus provides a shared and common understanding of a domain that can facilitate communication between users and applications.

In the framework proposed in [20], the MPEG-7 was extended with the domain ontology formalized using a logical formalism. In this system, syntactic data are presented in MPEG-7 standard and semantic data are described using the ontology's vocabulary. Ontology of artistic concepts was employed in [13], which included visual concepts at the intermediate level, and high-level concepts at the application level. Color and brushwork concepts were combined with low-level features and a transductive inference framework was used to annotate high-level concepts to the image blocks. A formalized core context-based multimedia ontology model was developed in [21] to facilitate multimedia semantic organization and management. A knowledge infrastructure and an experimentation platform for semantic annotation were presented in [17]. Here, ontology was extended and enriched to include low-level audiovisual features and descriptors. A multi-ontology based multimedia annotation model [1], [7] was proposed for multimedia access to address different users' requests.

All of the abovementioned systems used ontology to perform semantic indexing and annotations without using a Resource Description Framework (RDF) or Web Ontology Language (OWL) and reasoning techniques. RDF provides a means of adding semantics to a document. It is an infrastructure that encodes, exchanges, and reuses information on structured metadata. RDF allows multiple metadata schemas to be read by users and machines, and provides interoperability between applications. OWL is designed for use by applications that need to process the content of information instead of just presenting information to users. OWL facilitates greater machine interpretability of web content than that supported by XML and RDF by providing additional vocabulary along with formal semantics [21].

Some studies in multimedia retrieval, based on ontology, use RDF or OWL to describe semantic context [4], [5]. The DS-MIRF framework [3] was established to support interoperability of OWL with MPEG-7/21, so that domain and application ontology expressed in OWL could be transparently integrated with MPEG-7/21 metadata. This allowed applications that recognized and used constructions provided by MPEG-7 to make use of domain and application ontology. Retrieval performance was thus enhanced, and user interaction was provided with audiovisual content. In [14], OWL was not only used to describe semantic content, but also, by means of a reasoning method, inferred more complete queries. Their approaches also exploit the domain knowledge embedded into ontology to learn a set of rules for semantic video annotation.

C. Bayesian Inference for Ontology Matching

In order to match queries and the domain ontology, a Bayesian network [10] is used to calculate the similarities. The Bayesian network was used to associate low-level features with query concepts [8], thus decrease the gap between syntactic and semantic data. A system [2] using a Bayesian network and support vector machine (SVM) was developed to classify relevant documents by ontology concept, and perform semi-automatic annotations.

However, in order to calculate similarity using a Bayesian network, dependent ontology concepts (classes) are required. When converting ontology into BayesOWL type, there cannot be any equivalence between concepts; otherwise a loop in the calculation of conditional probabilities may occur. To reduce the complexity of a Bayesian network, ontology concepts are assumed independent of each other. Thus, naïve Bayes can be used in the ontology matching process. A naïve Bayesian classifier is deployed to calculate the similarity between query and domain ontology, where query is transferred into ontology. RDF triples are regarded as documents and classify the documents based on ontology.

III. SEMANTIC-BASED IMAGE RETRIEVAL FRAMEWORK

The proposed framework, as shown in Fig. 1, is a semantic-based image retrieval approach. The framework also supports low-level features image retrieval. The more details of the design are presented in next section, in which we explain the methodologies used in the implementation.

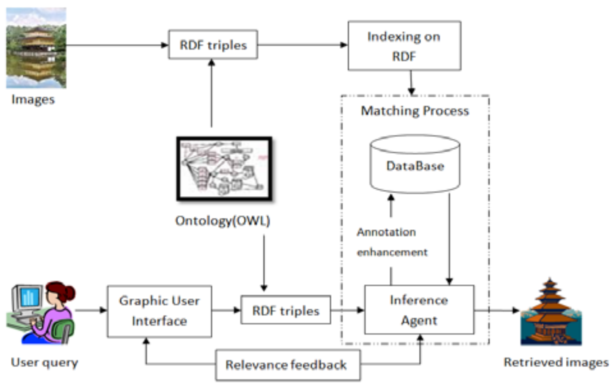

Fig. 1 Proposed framework for semantic-based image retrieval

The proposed framework consists of three major processes: RDF translation and indexing, user query, and matching process.

- The semantic annotations of images which are to be stored in the database are translated into RDF triples according to the domain ontology. Indexing is then implemented on these RDF triples.
- When image seekers enter queries through the graphical user interface, they can choose to use semantic query or an image query. For an image query, image seekers may

provide some preliminary semantic annotation manually, along with the image. For a semantic query, the query is translated into RDF triples. The RDF triples are then forwarded to an inference agent for matching images.

- In the matching process, the matched images are provided to the image seekers for feedback as to whether the selected images satisfy their needs or not. These images, which have strong relevance to the queries, are sent back to the inference agent for annotations enhancement, so that afterward other image seekers can retrieve images more precisely.

The data resources collected for our experiments are scenic images of various national parks in different countries, and in different seasons. Some of the images have already had semantic annotations. Based on the collected images, domain ontologies are constructed that appropriately describe the relationship between classes, or properties of classes. The MPEG-7 descriptors of the images are extracted as properties of classes. Hence, the descriptors can be simultaneously applied with the ontology in order to represent images. The MPEG-7 descriptors can also be used directly for content-based image retrieval.

Images can be annotated manually before translating them into RDF triples. The translation of the semantic annotated image is not only based on domain ontology, but also using MPEG-7 descriptors. Since ontology is built by including MPEG-7 descriptors as a property of class, the framework can retrieve images from the repository and enhance annotations of the images using their MPEG-7 descriptors. After images are translated into RDF triples, these images are indexed on the triples for improved image retrieval performance.

The functions of the inference agent are query reasoning, ontology matching, and image annotation enhancement. When an image seeker issues a query through the graphic user interface, one may choose to use an image query or a semantic query. For an image query, the query image's low-level features are extracted and used to match with those of images in the database. For a semantic query, the query is translated into RDF triples and the RDF triples are constructed into a small ontology. Then, the framework uses the small ontology to find matches within the domain ontology.

Before constructing a query into a small ontology, query reasoning must be firstly performed to derive more related semantic keywords. For example, while a semantic query "church near the sea" is entered from the graphic user interface, the framework searches for "church" or "sea" (e.g., class or instance) in the domain ontology. In the domain ontology, "church" is a subclass of the "building" class. Thus, a semantic query "building near the sea" is derived. Similarly, the framework searches for relevant properties of "church" and finds the semantic query "church has color gray" or "church has window". These derived semantic keywords (e.g., class, super class, or properties) are constructing into the small ontology. This allows the framework to find more relevant images in ontology based matches.

A naïve Bayesian classifier is employed to match domain ontology to the query ontology. The corresponding relevant images are provided to the image seekers for relevance feedback. Any negatively relevant image is discarded and positively relevant images are used for further processing. The positive relevant images are then used to search for images within the database using their MPEG-7 descriptors. The inference agent annotates the searched un-annotated images with keywords of the positive relevant images. To search annotated images, the inference agent enhances their annotations with more keywords. Hence, query performance can be improved in future queries.

IV. SYSTEM DESIGN AND IMPLEMENTATION

The details of the system design and methodologies used in the implementation are provided in this section. How to build domain ontology? We can use Lucene image retrieval (LIRE[2]) to implement low-level feature image queries, and use naïve Bayesian inference in the matching process are described.

A. Ontology Construction

The basic architecture of building ontology includes four parts: database storage, application interface (API), high-level ontology objects, and applications. MySQL is employed to store data. Storage API provides a single channel for ontology objects to access data. Jena[3] is deployed to modify classes, properties, or instances of ontology, and to query classes or instances on ontology. Neo4J[4] is used to traverse ontology. Both Jena and Neo4J provide efficient ontology queries on RDF and OWL. High-level ontology objects consist of ontology, concept, lexicon, terms, and context. High-level ontology objects are used to define and interpret class concept, role of class or property, rules between classes or subclasses, and context within ontology. Applications provide the user interface, for issuing queries and receiving query results.

The development tool Protégé[4] is used in building a new ontology. The new ontology construction process is as follows: confirm scope and domain, identify important terms for ontology, and define hierarchical of classes and the properties of classes. Some elements are used as the concepts and properties in ontology. For example, "content" is used to describe the scenario portrayed in the image. The "content" is divided into more detailed concepts, such as country, objects, scene, and events. "Multimedia feature" such as format, size, and capacity is used to describe more information about the images. Fig. 2 illustrates the partial structure of domain ontology. After defining the necessary terms for domain ontology, classes and their properties are identified and arranged in hierarchically class. For example, the class "objects" has properties "have color" and "have texture". Instances can be attached to ontology. For example, the class "country" has instances like Japan, Iran, and Italy. Other well constructed ontologies can be reused in building new domain ontology. LSCOM[5] defines many concepts (classes) which are helpful in building this landscape domain.

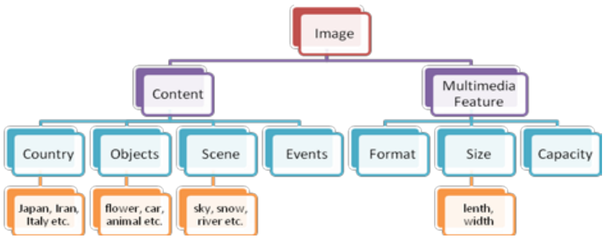

Fig. 2 Structure of a domain ontology

[2] http://www.semanticmetadata.net/lire/
[3] http://www.jena.sourceforge.net/
[4] http://protege.stanford.edu/
[5] http://www.lscom.org/?dc788400

B. Image Annotation and Translation into RDF Triples

PhotoStuff is employed to perform image annotations. First, the domain ontology is imported into PhotoStuff. Second, interesting areas of the image is circled for annotation. Third, the Instance Form is used to create a new instance with which the properties and annotations are filled in. Finally, the image annotations are translated into RDF triples. These steps furnish the manual annotation for an image. Jena is used to modify annotation or add new annotation to an image.

After image annotations are translated into RDF triples, an image index [12] [19] is built based on RDF triples with LARQ, which combines Lucene and SPARQL. LARQ provides methods for indexing RDF triples using SPARQL language when storing RDF triples in the database. Since RDF triples are composed of <S, P, O>, which represent subject, predicate and object, reasoning can be performed on RDF triples. In addition, naïve Bayes is used to calculate the similarities of the query ontology and domain ontology using their RDF triples. Jena is open source software and has large numbers of API for RDF and RDFS' data retrieval. It provides ways to express objects, such as graphs, resources, properties, and literals. Jena with SPARQL language is used to access classes, properties, and instances of ontology.

C. Image Retrieval with Low-level Features

The framework provides image retrieval using low-level features. The low-level features are represented using MPEG-7 descriptors, which can be used for image queries. LIRE software package is employed in the retrieval of images with color and edge low-level features. The LIRE is used to implement the extraction of feature vectors from images and convert the feature vectors into text data type. The framework then indexes the feature vectors of images as documents and stores them in a database. The image indexing procedure of LIRE is similar to the Lucene documents indexing procedure.

D. Naive Bayesian Classifier

In the proposed framework, image annotations are translated into RDF triples and indexed by LARQ. Semantic query is also translated into RDF triples and constructed into the query ontology. The naïve Bayesian classifier is used to match the query ontology with the domain ontology. To calculate the similarity between query ontology and domain ontology, Neo4J is applied to traverse domain ontology and to search for the context (classes or their super classes) and properties of the query terms.

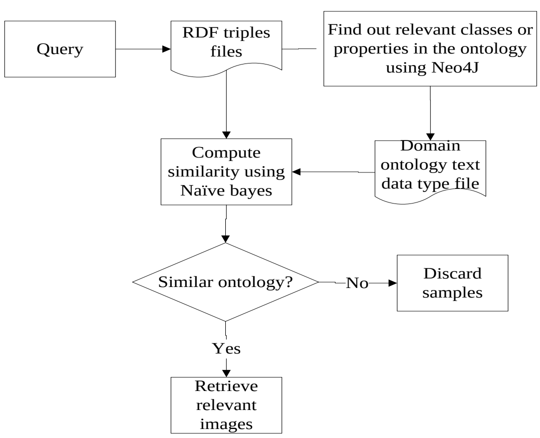

Fig. 3 Similarity computation of the query and domain ontology

For example, in order to find the context for "apple", Neo4J first finds the nodes with term "apple" from the domain ontology. Neo4J then searches for the "apple" node's parent node to find content about "apple", and its children nodes to retrieve the "apple" node's properties. The relevant classes and properties with query terms are converted into text data type and stored in the database. Thus, naïve Bayes can be employed to compute similarities of the query ontology with relevant domain ontology. The query ontology is classified as relevant or irrelevant with to the domain ontology. The above procedures are summarized in Fig. 3.

The naïve Bayesian classification procedures are as follows.

(1) We assume the query terms are n keywords denoted as: $X = (x_1, x_2, \cdots, x_n)$;

(2) We suppose that there are m classes which represent similar domain ontology $C_1, C_2, \ldots, C_m$;

(3) The maximum probability of X belongs to the most probable class C_i is estimated. The naïve Bayesian classifier with the maximum a posteriori (MAP) decision rule is defined as:

$$\arg\max_{C_i} P(C = C_i) P(X \mid C = C_i).$$

The naïve Bayesian classifier assumes class conditional independent, hence, the MAP decision rule can be expressed as:

$\arg\max_{C_i} P(C = C_i) \prod_{k=1}^{n} P(X = x_k \mid C = C_i)$ where x_k are keywords of the query.

(4) Keywords are augmented with weights. The term $P(X = x_k \mid C = C_i)$ is replaced by $P(X = w_k x_k \mid C = C_i)$.

To train the naïve Bayesian classifier, there must have some training annotated samples. For example, a semantic query is "church near the sea" and there are 100 RDF training samples in the database. Each RDF sample represents an image's annotation. Now, the query "church near the sea" is expressed by $X = (church, sea)$. Assume the database has been searched for similar RDF samples, and obtained 5 relevant RDF samples with ontology C_1, and 10 relevant RDF samples with ontology C_2. In the query ontology, "church" appears three times and "sea" appears twice. Within 100 RDF training samples, "church" appears 25 times and "sea" appears 20 times. The weight of the query terms "church" and "sea" is calculated similar to TF-IDF [22].

The weight of "church": $w_1 x_1 = 3 * \log(100/25) = 1.8$

The weight of "sea": $w_2 x_2 = 2 * \log(100/20) = 1.4$

$$P(church \mid C_1) = \frac{1.8}{5} = 0.36$$

$$P(sea \mid C_1) = \frac{1.4}{5} = 0.28$$

$$P(X \mid C_1) = \prod_{k=1}^{2} P(w_k x_k \mid C_1) = 0.36 * 0.28 = 0.1008$$

$$P(C = C_1)P(X \mid C_1) = \frac{5}{10+5} * 0.1008 = 0.0336$$

$$P(church \mid C_2) = \frac{1.8}{10} = 0.18$$

$$P(sea \mid C_2) = \frac{1.4}{10} = 0.14$$

$$P(X \mid C_2) = \prod_{k=1}^{2} P(w_k x_k \mid C_2) = 0.18 * 0.14 = 0.025$$

$$P(C = C_2)P(X \mid C_2) = \frac{10}{10+5} * 0.025 = 0.01666$$

According to the above calculation, the value of $P(C = C_1)P(X \mid C_1)$ is larger than the value of $P(C = C_2)P(X \mid C_2)$. Thus, the RDF samples with ontology C_1 are the most similar to the query.

V. EXPERIMENTAL DESIGN AND RESULTS

This study focuses on developing a framework for image retrieval based on ontology and demonstrating the validity of the proposed matching process. The 500 scenic images extracted from the image database [6] are used in our experiments. Forty percent of the extracted images were originally annotated.

A. Experimental Design

To demonstrate our framework's feasibility and improvement in image retrieval performance, several new ontologies are built, some related to the scenic images and some unrelated to the domain. After building the ontology, some of the images are manually annotated based on classes or properties of the ontology. The annotations are then translated into RDF samples which are in term used in the matching process. During experiments, more ontologies may be built or more annotations may be added manually to images. The more ontologies built, the more relative images are retrieved. However, those ontologies whose classes or properties are not used in annotations are not helpful for query.

In the naïve Bayesian inference process, a threshold is chosen for estimating the likelihood probability, $P(X = w_k x_k \mid C = C_i)$, in order to decide whether domain ontology is matched with the query ontology. In the initial experiments, due to the limited number of annotated images, a small threshold value should be chosen, increasing it gradually as the number of annotated images increased in the successive experiments. Table I provides experimental results regarding the selection of different thresholds. Some ontologies are built during experiments in order to test whether the similar ontology can be matched. There are 2, 9, and 15 domain ontologies built for the initial query, six queries, and ten queries, respectively. The results show that the number of relevant and irrelevant ontologies matches the different thresholds. With a higher threshold, less ontologies are matched because of the limitation of RDF triples (annotated images). After some experiments have been done, the likelihood probabilities of annotated keywords are increased, since the weights of query keywords are increased as more images are annotated and more corresponding RDF triples are added into database.

TABLE I NUMBER OF MATCHED ONTOLOGIES WITH DIFFERENT THRESHOLD

Experiments	**Thresholds / relevant or irrelevant ontology matched**			
Initial query	0.005	0.01	0.02	0.05
	Relevant: 1 Irrelevant: 1	Relevant: 1 Irrelevant: 0	Relevant: 0 Irrelevant: 0	Relevant: 0 Irrelevant: 0
Six queries	0.01	0.05	0.07	0.10
	Relevant: 7 Irrelevant: 2	Relevant: 7 Irrelevant: 0	Relevant: 4 Irrelevant: 0	Relevant: 0 Irrelevant: 0
Ten queries	0.1	0.2	0.3	0.4
	Relevant: 12 Irrelevant: 3	Relevant: 12 Irrelevant: 0	Relevant: 6 Irrelevant: 0	Relevant: 2 Irrelevant: 0

Since the threshold is selected to decide whether the ontology is similar to the query ontology, multiple similar ontologies are matched on some occasions. For example, there are two ontologies: one is relevant to buildings; the other is relevant to landscapes which have building class. If their likelihood probabilities of similarity are higher than the threshold, then we use both the buildings and landscapes ontologies to select similar images, i.e., the query ontology is classified into both the buildings and landscapes ontologies.

Precision and recall [6] are used to evaluate the experimental results. Precision is the percentage of retrieved documents which are relevant to the query. Recall is the percentage of documents that are relevant to the query and are retrieved. The framework allows image seekers to mark which images are relevant to their queries. The image seekers may answer yes or no. Relevance feedback returns to the inference agent.

B. Experimental Results

In this subsection, experimental results are presented and compared with the performance of the naïve Bayesian inference and with that of the smart ontology mapping (SOM[7]) algorithm.

1) *Precision and Recall of Using Consecutive same Queries:*

In order to evaluate the query's effect on precision and recall, precision and recall are measured on six different semantic queries. Each semantic query is issued four times consecutively in the experiment. The six semantic queries used are "there have temple and tree", "there is a kola", "church near the sea", "there have sky", "there is a lake", and "there have stone". The trend of precision and recall for the six semantic queries are shown in Fig. 4 and Fig. 5. Note that in the experiments threshold is selected as 0.05 for the naïve Bayesian inference.

From the results, the semantic query "there have temple and tree" enhances annotations effectively by using low-level features of temples in the first query. The precision is raised dramatically in the second query, because the most relevant images are now annotated. The precision is not much

[6] http://www.cs.washington.edu/research/imagedatabase/

[7] http:/www.som.sourceforge.net/

improved in the third and fourth queries. The trend is the same for recall.

For a semantic query "there is a koala", there are some images with a koala in a tree found, but it is ineffective in annotation to use low-level features of trees in the first query.

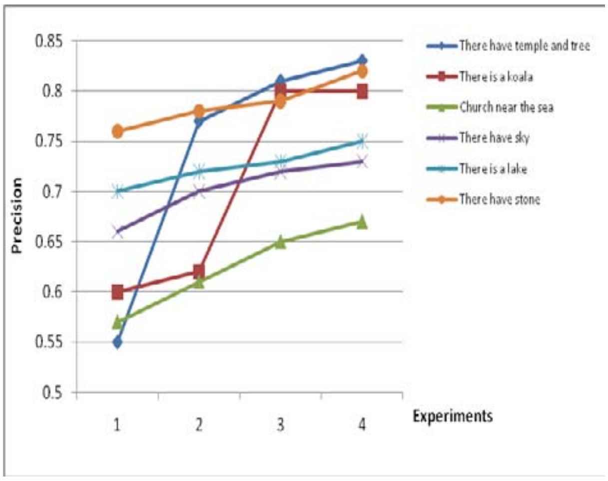

Fig. 4 Precision of the six semantic queries

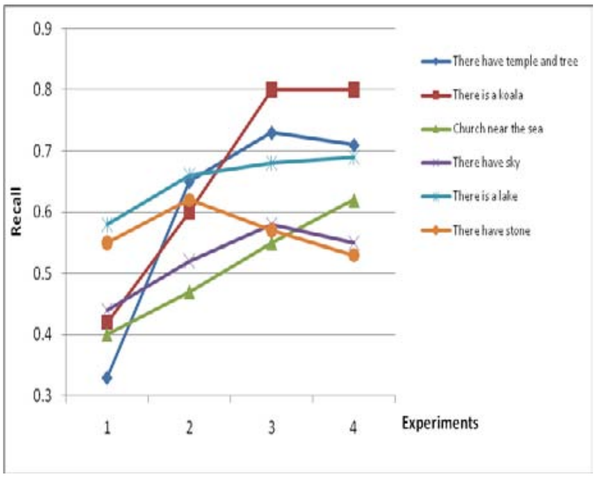

Fig. 5 Recall of the six semantic queries

The keyword "koala" is annotated on some of the koala in tree images or tree images, where tree images belong to unrelated ontologies. Thus, the precision is not improved in the second query since not many images are annotation enhanced. There are some koala images without trees found in the second query, and their annotations are enhanced. This results in improving the precision in the third query. In the second query, the weights of RDF samples are calculated through naïve Bayesian inference and these tree images without koalas are not selected in the inference process. This is because there are other annotations in the tree images without koalas and the weights of "koala" in the images are low. The recall improves in the second and third queries since the number of relevant koala images is small in the database.

For the other semantic queries, there are initially many images with annotations. Their precisions are high in the first query and improve gradually. Most of the recalls decline as relevant images increase due to enhanced annotations. When numbers of annotated images increase, the relevant images in the database also increase. Thus recall decline.

2) Precision and Recall of Using Mixed Queries:

The precision and recall of experiments are evaluated with mixed semantic queries. Fig. 6 shows the precision and recall of the twelve semantic queries.

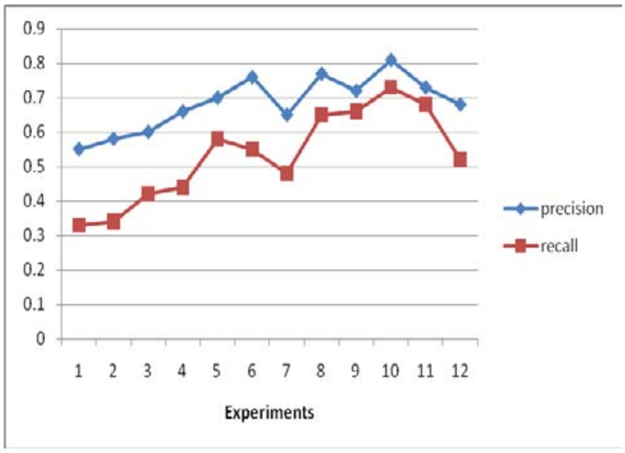

Fig. 6 Precision and recall of mixed semantic queries

The twelve semantic queries used are "there have temple and tree", "church near the mountain", "church near the sea", "there have sky", "there is a lake", "there have stone", "church near the sea", "there have temple and tree", "there is a lake", "there have temple and tree", "church near the sea", and "there is a lake". Note that we intentionally repeat the last six semantic queries with the previous queries. The use of the previous issued semantic queries is to measure whether the precision and recall are increasing. For example, the Experiment 3, 7 and 11 are the same semantic query "church near the sea". The precision results are, respectively, 0.6, 0.65 and 0.73 and the recall ratios are, respectively, 0.42, 0.48 and 0.68. Because the number of annotated images is increased since the first time the semantic query was issued, the chance to retrieve images with the enhanced keywords is also increased when re-issuing the query. Since more images are annotated with enhanced keywords, the precision and recall are improved after more queries are issued.

Fig. 7 provides the number of the original annotations and enhanced annotations of retrieved images in the experiments. The improvement in precision and recall can be observed from the number of enhanced annotaions. For example, in Experiment 1, there are six enhanced annotations. In the following, Experiment 8 and 10, both precision and recall are improved for the query "there have temple and tree", since the keywords "temple" and "tree" are annotated to some images. This means that, for the next query, more relevant images can be retrieved.

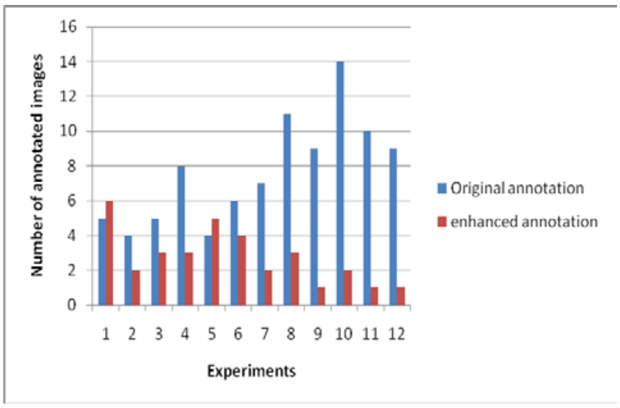

Fig. 7 Number of original annotations and enhanced annotations

3) Comparison of the Naïve Bayesian Classifier and SOM: The performance of the naïve Bayesian classifier is compared with SOM. Since SOM is used to perform ontology mapping and merging, it can only find the most similar ontology at given time. The naïve Bayesian classifier can find multiple similar images according to the user's semantic query. Fig. 8 shows the result of the semantic query "church near the sea" using the naïve Bayesian classifier and SOM algorithm. From the results, the SOM first find church images without having sea. But the naïve Bayesian classifier find images with both church and sea simultaneously. Fig. 9 is the query results of "there have temple and tree". Although the SOM algorithm can find images with a temple and a tree, it still finds images which have no tree. However, the naïve Bayesian classifier finds more numerous relevant images with both temple and tree. Thus, from the compared results, the proposed method retrieves more precise query results.

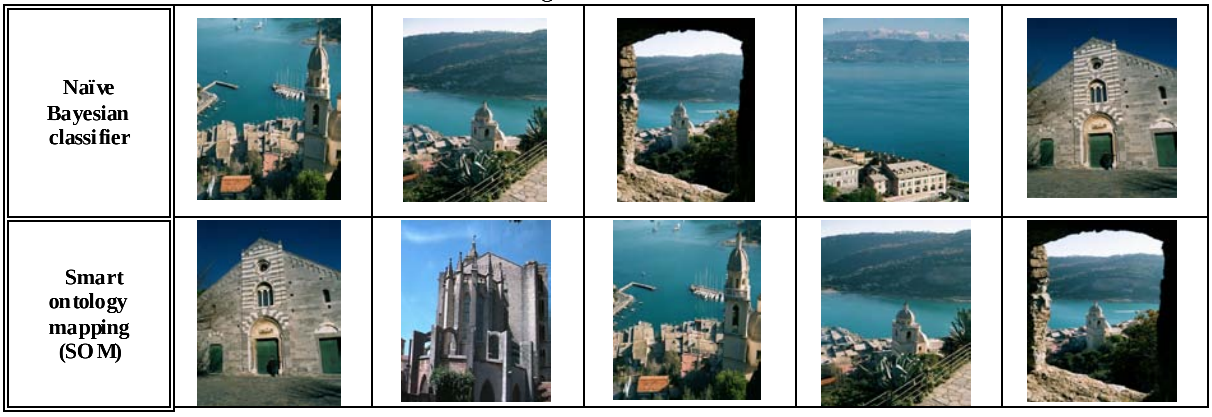

Fig. 8 Query results of "church near the sea"

Fig. 9 Query results of "there have temple and tree"

VI. CONCLUSIONS

In experiments, the proposed semantic image retrieval framework has shown that the approach notably improves the precision of image retrieval by incrementally enhancing image annotations through relevance feedback. With the domain ontology, the framework provides a scheme which facilitates the sharing of ontology information among image seekers. As shown in Fig. 4 and Fig. 6, the methods suggested for improving image retrieval precision demonstrably improve search results.

The experiments have shown that the proposed framework finds more numerous similar images through domain ontology, while image seekers can gradually add more ontology into the framework, and the naïve Bayesian classifier does indeed find all relevant ontology to the query. Thus, more images are retrieved for relevance feedback. Furthermore, the framework uses relevant feedback images from users to search for more numerous similar images by using their low-level features from annotated or un-annotated images. This process enhances the similar semantic images with more annotations.

From the experiments, constructing more detail in ontology affects query precision and recall. By defining more properties of ontology classes (concepts), more annotations can be enhanced to the similar images. For example, for a class "fruit" and its properties "name", the ontology can only make instances (annotations) with fruit name. If more properties for the class, such "colour" and "shape" are defined, then more instances can be annotated about fruit with colour and shape. When making queries, with more annotations, more precise semantic images are able to be retrieved, thus increase the recall and precision. Since building ontology is subjective, this also affects the query results. If well constructed or formal ontology is available for use, the framework can retrieve more precise results.

REFERENCES

[1] D. Aijuan and H. Li, "Multi-ontology Based Multimedia Annotation for Domain-specific Information Retrieval", *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, vol.2, pp.158-165, 2006.

[2] E. Chang, K. Goh, and G. Wu, "CBSA: Content-based Soft Annotation for Multimodal Image Retrieval Using Bayes Point Machines", *IEEE*

Transactions on Circuitsand Systems for Video Technology, vol.13, pp.26–38, 2003.

[3] T. Chrisa, P. Panagiotis, and C. Stavros, "Interoperability Support between MPEG 7/21 and OWL in DS-MIRF", *IEEE Transactions on Knowledge and Data Engineering, vol. 19(2)*, pp.219-232, 2007.

[4] T. Chrisa, P. Panagiotis, and C. Stavros, "Integration of OWL Ontologies in MPEG-7 and TV-Anytime Compliant Semantic Indexing", *Lecture Notes in Computer Science*, vol.3084, pp.143-161, 2004.

[5] T. Chrisa and C. Stavrous, "A Multimedia User Preference Model that Supports Semantics and Its Application to MPEG 7/21", *Proceedings of Multi-Media Modelling*, 2006.

[6] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves", *Proceedings of ICML*, pp.233-240, 2006.

[7] J. M. Gomez-Berbis, R. Colomo-Palacios, J. Lopez-Cuadrado, I. Gonzalez-Carrasco, and A. Garcia-Crespo, "SEAN: Multi-ontology semantic annotation for highly accurate closed domains". *International Journal of the physical Sciences*, vol. 6(6), pp.1440-1451, 2011.

[8] E. Izquierdo and Q. Zhang, "A Bayesian Network Approach to Multi-feature Based Image Retrieval", *Lecture Notes in Computer Science*, vol. 4306, pp.138–147, 2006.

[9] N. James, K. Todorov, and C. Hudelot, "Ontology Matching for the Semantic Annotation of Images", *IEEE International Conference on Fuzzy Systems*, 2010.

[10] K. Jayech and M. Mahjoub, "New approach using Bayesian Network to improve content based image classification systems", *International Journal of Computer Science Issues*, vol. 7, issue 6, pp. 53-62, 2010.

[11] P. J. Kuo, A. Terumasa, and H. Yasuda, "MPEG-7 Based Dozen Dimensional Digital Content Architecture for Semantic Image Retrieval Services", *Proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service*, pp.517-524, 2004.

[12] X. Y. Lee, J. S. Yuan, and X. M. Yang, "Intelligent Search Engine for XML Based on Index and Domain Ontology", *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics*, pp.4501-4506, 2006.

[13] L. Liza, T. S. Chua, and J. Ramesh, "Annotation of Paintings with High-level Semantic Concepts using Transductive Inference and Ontology-based Concept", *Proceedings of the 15th international conference on Multimedia*, pp.453-452, 2007.

[14] B. Marco, D. B. Alberto, and S. Giuseppe, "Learning Ontology Rules for Semantic Video Annotation", *Proceeding of the 2nd ACM workshop on Multimedia semantics*, pp.1-8, 2008.

[15] F. Nastaran, "MPEG-7 in Practice: Analysis of a Television News Retrieval Application", *Journal of the American society for information science and technology*, vol.58, pp.1364–1366, 2007.

[16] W. Otto and C. Oscar, "An MPEG-7 Database System and Application for Content-based Management and Retrieval of Music", *5th International Conference on Music Information Retrieval*, 2004.

[17] K. Petridis, S. Bloehdorn, and C. Saathoff, "Knowledge representation and semantic annotation of multimedia content", *IEE Proceedings on Visual Image Signal Process*, vol. 153(3), pp.255-262, 2006.

[18] D. Ramprasath and R. M. Kamesh, "Compact Combination of MPEG-7 Color and Texture Descriptors for Image Retrieval", *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol.1, pp.387-391, 2004.

[19] M. Roshan, Y. K. Lee, and S. Lee, "Exploiting Path Indexes to Answer Complex Queries in Ontology Repository", *International Conference on Computational Science and Its Applications*, pp.56-61, 2009.

[20] H. Smira, B. Salima, and V. Athena, "A Logic Based Approach for the Multimedia Data Representation and Retrieval", *IEEE International Symposium on Multimedia*, 2005.

[21] I. Tanko and E. Emmanuel, "Context-Based Multimedia Ontology Model", *Third International Conference on Autonomic and Autonomous Systems*, pp.2, 2007.

[22] H. J. Zeng, Q.C. He, Z. Chen, W. Y. Ma, and J. Ma, "Learning to cluster web search results", *Proceedings of the 27th annual international ACM SIGIR conference on Research and Development in Information retrieval*, pp.210-217, 2004.